

Paolo Labinaz

POLARIZATION AND VIOLENT COMMUNICATION: A SPEECH-ACT THEORETICAL PERSPECTIVE

Introduction

This paper deals with polarization and violent communication in online discussions from a speech-act theoretical perspective. While socio-epistemic concepts like echo chambers and epistemic bubbles are frequently used to analyze ideological conflicts, less attention has been given to how language shapes group identity and fosters hostility in polarized digital spaces.

In these environments, polarization and violent language are closely intertwined. Violent language, undeniably common on social media platforms, is often both a symptom and a driver of polarization. Scholars addressing this issue typically focus on its emotional, cognitive, and behavioral effects (such as manipulation or psychological harm), which naturally leads to an emphasis on the perlocutionary dimension. However, to truly understand polarization, we also need to look at the illocutionary dimension.

My aim is to show how communicative strategies aimed at discrediting and delegitimizing opponents in online discussions can exploit the standard conversational dynamics of illocutionary acts, particularly assertions, to amplify polarization and reinforce echo chambers, ultimately fueling violent communication. I will examine the distortion of these dynamics in polarized environments by first drawing on an Austin-based framework and then illustrating them with examples from social media discussions.

The Austinian framework

To understand how language shapes group identity and fosters hostility in polarized online spaces, I will start by examining the dynamics of illocutionary acts, that is, what we do in uttering a certain sentence, provided certain conditions are met. This will be done through an Austinian framework, as elaborated by Marina Sbisà (1984; 2007; 2023, 1–4; see also Caponetto & Labinaz, 2023, 1–17).

According to Austin (1975, 14–15, 116–118), an illocutionary act succeeds when it follows an “accepted conventional procedure” comprising three key elements:

- initial conditions: the speaker (and sometimes the listener) must be in the appropriate position, and the procedure must be invoked under suitable circumstances.
- procedural steps: the speaker must utter a sentence with recognizable linguistic features.
- designed effects: if these requirements are met, the procedure achieves its expected effects through intersubjective agreement, producing changes in the rights and obligations of both speaker and audience, as Sbisà (2023, 2) notes.

Illocutionary acts vary based on the normative changes they introduce, which establish legitimate expectations for subsequent exchanges, thereby defining the standard conversational dynamics of a given illocutionary act. For example, by uttering “Shut up!”, a speaker typically invokes the procedure for issuing an order. If she does so from a recognized position of authority, she would be typically taken as issuing an order, which obligates the listener to be quiet and grants her the right to enforce compliance. Thus, the issuance of an order creates legitimate expectations that actions aligned with its normative effects should follow.

Challenges in online communication

While originally developed for face-to-face verbal interactions, this speech act framework encounters significant challenges when applied to online contexts. Here, I will highlight three main obstacles, each connected to a fundamental aspect of the illocutionary procedure introduced earlier:

- **limited information:** the scarcity or unreliability of information about interlocutors makes it difficult to assess whether the initial conditions are met. Participants often rely on alternative signals, such as self-proclaimed competence, writing style, profile pictures, usernames, posting history, and follower counts, to gauge their own and others' credibility (see, e.g. Labinaz & Sbisà, 2017). However, these alternative signals can be easily manipulated, thereby allowing individuals to create a false or misleading image of themselves. This can result in misattributing illocutionary force to utterances when the speaker's actual conditions do not justify it or failing to recognize illocutionary force where it would otherwise be justified, due to mistrust.
- **lack of non-verbal cues:** in face-to-face communication, non-verbal and paralinguistic cues reinforce linguistic force indicators. Online, these cues are absent, making illocutionary acts harder to recognize. Alternatives like all-caps or emoticons convey tone but can be ambiguous, increasing the risk of misinterpretation (see, e.g., Dresner & Herring, 2010);
- **ambiguous audience boundaries:** a single post often addresses diverse audiences, complicating the predictability of how illocutionary force will be received. Asynchronous communication makes this problem worse, as the delay in feedback leaves misunderstandings unresolved for longer, disrupting expectations regarding conversational dynamics.

Although these challenges do not undermine the relevance of speech-act theory for online discussions, they emphasize the necessity of addressing the distinct characteristics of digital communication when examining the performance of illocutionary acts in these contexts (see, e.g., Labinaz & Sbisà, 2021a, 68–69; 2021b, 148–153).

Conversational dynamics in polarized discussions

Online platforms are key areas for public discussions, where assertions often initiate interactions. When social networks users discuss contentious topics like immigration, healthcare, or climate change, conversations often heat up, leading to a wave of comments. In non-polarized discussions, disagreements typically lead to responses examining an assertion's validity, requesting evidence, or presenting counterarguments, enabling an evaluation of the assertion's truth and basis. As we will see, the same cannot be said in polarized contexts.

The illocutionary procedure for assertions

By relying on the Austinian framework, let us first consider the key elements that characterize the illocutionary procedure for assertions and explore how its effects establish legitimate expectations for subsequent moves (Labinaz, 2019, 104–124; Sbisà, 2020):

- initial conditions: the speaker must be recognized as having epistemic authority, which grants credibility to their claims. This entitles interlocutors to demand evidence or challenge competence;
- procedural steps: the required linguistic form is typically a straightforward, indicative sentence without modal verbs;
- designed effects: if the speaker is recognized as knowledgeable, the audience typically feel entitled to rely on the asserted content in their subsequent behavior, such as re-asserting it, using it as a premise in reasoning, or basing decisions on it.

These effects, along with the right to challenge the assertion, establish legitimate expectations after an assertion is made, thereby defining its standard conversational dynamics.

The standard dynamics of assertion in polarized contexts

Problems typically arise when people with ideologically divergent beliefs engage in public online discussions. In such contexts, they rarely act as mere individual agents. Instead, they interact from within what C. Thy Nguyen (2020) describes as “epistemic bubbles” or, more critically, “echo chambers”.

Epistemic bubbles are closed systems where participants access only a narrow range of information. Within these spaces, “bootstrapped corroboration” occurs, where repeated exposure to like-minded perspectives inflates confidence in beliefs, potentially distorting reality (Thy Nguyen, 2020, 143–145).

Echo chambers are even more harmful, as members not only access a limited set of information but also actively discredit and dismiss outsiders’ views, not based on rational evaluation, but simply because they come from outside the group (Thy Nguyen, 2020, 145–150).

In echo chambers, assertions often serve as markers of ideological stance and group loyalty, clearly dividing “us” from “them”. Within the group, such assertions are readily accepted, but they often provoke hostility from outsiders. Instead of encouraging dialogue, assertions end up signaling membership, undermining their standard conversational dynamics.

Exploitation of conversational dynamics in polarized spaces: a case study

A recent case study can help illustrate how discussions on even seemingly trivial topics can escalate polarization, thereby reinforce group identities and fueling hostility.

On June 1, 2024, Piazza del Popolo in Rome hosted a rally concluding the electoral campaign for the governing party Fratelli d’Italia, featuring a speech by Prime Minister Giorgia Meloni. The online newspaper *Fanpage* posted a photo hinting at low attendance, which quickly ignited heated debates

between her supporters and critics¹. Both sides began sharing photos and videos to support their claims, with one asserting that the square was half-empty and the other insisting it was full.

To illustrate the typical ideological maneuvering that occurs in discussions of this nature, I will examine a series of examples taken from a collection of comments on posts addressing this debate on the social network X (see Appendix).² Specifically, I will classify these examples into four key communicative strategies that, as I will suggest, exploit the dynamics of assertion to discredit opponents and their sources, often escalating to extreme linguistic violence.

Beyond disagreement and rejection

In polarized online debates, conversational dynamics escalate beyond a disagreement or rejection of opposing claims. Though some users still try to refute their opponent's assertion with counterevidence (e.g., personal evidence, as illustrated in example 1 or sound reasoning, as shown in example (2), such cases are growing increasingly rare:

(1) T1

Come fa a dire che è falsa lei c'era? Io sì ed era esattamente così la piazza
 'How can you say it's fake? Were you there? I was, and the square was exactly like that'

(2) T2

Ha sovrapposto l'audio a immagini antecedenti. Lo si può fare con qualsiasi programma che trovi GRATIS su internet
 'They overlaid the audio on earlier images. You can do it with any program you find FREE on the internet'

.....
 1 *Fanpage* (2024): Meloni non riempie la piazza: il comizio decisivo di piazza del Popolo è un mezzo flop. <https://x.com/fanpage/status/1796927183153242130> (last access: 2024-11-27).

2 For privacy reasons, all personal information about the X users involved in the analyzed discussions (including their names, surnames, and nicknames) has been removed. Each analyzed example is preceded by an acronym (e.g., T1, T2, etc.), and the Appendix provides a link for each post, along with details about its popularity (views, comments, reposts, and likes). All posts were published on June 1, 2024, the same day the rally took place.

However, in many cases, these argumentative moves are overshadowed by more hostile dynamics, as disagreements escalate into accusations of dishonesty or deliberate fabrication, reframing evidence not just as false but as an intentional act of deception:

(3) T3

sai sparare solo cazzate

‘you only know how to spout bullshit’

(4) T1

Si vede lontano un miglio che è tarocca. L'istituto luce era più capace

‘You can see from a mile away that it’s fake. The “Istituto Luce” [Fascist film corporation] was more capable’

(5) T3

Foto falsa. Siete ridotti male avete paura. Ma mentite a voi. Non a noi. [...]

‘Fake photo. You’re in bad shape, you’re scared. But you’re lying to yourselves [...]

These accusations are often accompanied by mocking or aggressive language:

(6) T1

Ah, l’ho fatto io il check. È fake. Che figura di merda

‘Ah, I checked it myself. It’s fake. What a shitty mess’

(7) T4

Se riprendi solo la metà [piazza] minchione 😏

‘If you only film the half [square] that’s full, you idiot’ 😏

(8) T3

Ma veramente sto ritardati che pensano di votare a sinistra (dato che non esiste più da tempo) fanno girare foto false e poi ne danno la colpa agli altri 🧐🤔?

‘But seriously, these idiots who think of voting for the left (which hasn’t existed for a long time) spread fake photos and then blame others for it’ 🧐🤔?

Even when users attempt to fact-check claims, as in 6, the discussion quickly turns into a personal attack aimed at discrediting opponents and calling their legitimacy into question using *ad hominem* arguments. Ridicule and polarization are further fueled by teasing tones and emoticons, as in 7 and 8. In a similar vein, accusations of deliberate fabrication, as in 6 and 8, shift the debate from evaluating evidence to attacking personal integrity, leading to the next strategy.

Beyond challenging one's competence and informative sources

Challenging an opponent's competence or expertise is a valid argumentative move. However, the following examples exceed that, also targeting people's mental capacity and integrity:

(9) T1

Peccato che non lo sia e dimostrate quanto il sinistro woke sia incapace non solo di pensare ma anche di verificare... bla bla bla

'Too bad it isn't, and you show how the woke left can't think or verify anything... blah blah blah'

(10) T3

Ma voi una verità la sapete dire sapete solo prendere bufale e spacciarle per verità ecco perché certi partiti esistono ancora per gente come voi

'But do you ever tell the truth? You only know how to take hoaxes and pass them off as truth, that's why certain parties still exist because of people like you'

(11) T3

Non è in malafede, semplicemente non ci arriva

'It's not bad faith, she just doesn't get it'

These comments suggest that opponents lack critical thinking skills (9), spread hoaxes (10), or lack basic understanding (11). Instead of refuting the arguments, they attack the intellectual credibility of the opponent's character. In 9, the tone mocks not just the individual but her broader ideology, deepening

polarization. Example 11 dismisses the opponent and her argument without any attempt at a priori legitimization of her position.

The attacks also extend to the discrediting of informational sources:

(12) T3

Fanpage... bisogna un po' rivedere la definizione di giornale qua direi

'Fanpage... we might need to revisit the definition of a newspaper here, I would say'

(13) T3

Mo' definire fanpage un giornale...mi sembra eccessivo

'Now, calling Fanpage a newspaper... seems a bit much to me'

(14) T3

Fatto e Fanpage, dico solo che schifo, siamo arrivati alla mistificazione della realtà, stile Putin

'Fatto and Fanpage, I just say what a disgrace, we've reached reality distortion, Putin-style'

These comments dismiss outlets like *Fanpage* and *Fatto Quotidiano* as illegitimate. Sarcasm in 12 and 13 questions whether *Fanpage* qualifies as a credible newspaper, while 14 escalates the critique, equating these sources to propaganda.

Silencing dissent

The following examples escalate aggression further, aiming to silence opposing voices entirely:

(15) T1

Shut up 🙄 e continua a prendere a piene mani da Fanpage siete bellissimi nella vostra ottusa incapacità

'Shut up 🙄 and keep taking everything from Fanpage, you're beautiful in your obtuse incapacity'

(16) T3

Ma vedi di non dire stupidaggini!

‘Stop saying stupid things!’

(17) T5

Zitta e prenditi un maalox!

‘Shut up and take a Maalox! [antacid]’

These comments issue straightforward orders, such as “shut up” or “stop saying stupid things”, serving as exercitives; their purpose is to limit the other party’s ability to participate. This kind of language is a claim of authority to determine over who may or may not contribute to the discussion.

These moves, rather than expressing disagreement, aim at the conversation’s dominance through the silencing of voices. This reinforces power dynamics: one side is using its power to control the discourse in order to marginalize the other perspective.

Discrediting opponents by using slurs and insults

The most extreme form of silencing in polarized discussions involves using slurs and insults to discredit opponents completely, such as in the following cases:

(18) T5

Zitto zecca con l’handicap

‘Shut up, you disabled tick’

(19) T1

Ritardato, tu un cervello non puoi comprarlo però. Sempre ritardato preso per il culo resterai

‘Retard, you can’t buy a brain though. You will always remain a retard and be mocked’

Comments like “you disabled tick” or “Retard, you can’t buy a brain” function as verdictives, categorizing opponents as inherently illegitimate due to certain characteristics attributed to them. In example 18, the term “disabled” targets

an individual's personal abilities, while “tick” (original in Italian: zecca) is a pejorative used to insult people with far-left ideologies. Example 19 escalates further, using ableist insults, labeling the opponent as a “retard” and asserting her permanent inadequacy. Such comments go beyond denying arguments and reach toward undermining the intellectual and social worth of a person.

Distorted conversational dynamics and the escalation of polarization

The strategies discussed exploit and distorted the conversational dynamics of assertions, undermining genuine discussion. The characteristics of online communication outlined in Section 3 (limited information, lack of non-verbal cues, and ambiguous audience boundaries) significantly amplify these distortions and reinforce echo chambers.

Limited information about interlocutors contributes to situations where other's credibility is constantly questioned, encouraging aggressive language and personal attacks to take control of the discussion. The absence of non-verbal cues complicates the recognition of illocutionary force, leading to frequent misinterpretations (such as reading sarcasm where sincerity is intended) which escalate tensions in already polarized contexts. Ambiguous audience boundaries can make conflicts worse, as assertions acceptable within echo chambers often provoke hostility when perceived as provocative or misleading by outsiders.

Within this online environment, the distortions of the standard conversational dynamics of assertions fuel polarization by prioritizing group loyalty over substantive discussion, framing opposing views as inherently illegitimate and fueling an “us vs. them” mentality. Such dynamics are reinforced through speech acts that claim authority over the legitimacy of participants and their perspectives. The combination of all these elements creates conditions for violent communication, making derogatory language more common and deepening hostility between groups.

Conclusion

This paper explored how polarization and violent communication in on-line spaces can be approached from a speech-act theoretical perspective. Understanding how the communicative strategies aimed at discrediting and delegitimizing opponents in online discussions interact with and distort the usual dynamics of illocutionary acts is key to uncovering the mechanisms that drive polarization and reinforce echo chambers. If we want to foster healthier online discussions, we need to recognize these distorted dynamics and take steps to counteract them. This would involve, but is not limited to, encouraging respectful disagreement, fostering engagement with diverse perspectives, and developing features on platforms that help mitigate polarizing and hostile speech.

Sources and literature

Austin, J. L. (1975): *How to Do Things with Words*. 2nd edition. Cambridge (MA), Harvard University Press.

Caponetto, L. & P. Labinaz (2023): Marina Sbisà's Deontic Approach to Speech Actions. In: Caponetto, L. & P. Labinaz (eds.): *Sbisà on Speech as Action*. Cham, Palgrave Macmillan, 1–26 DOI: https://doi.org/10.1007/978-3-031-22528-4_1

Dresner, E. & S. C. Herring (2010): Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force. *Communication Theory*, 20, 3, 249–268 DOI: <https://doi.org/10.1111/j.1468-2885.2010.01362.x>.

Labinaz, P. (2019): *L'asserzione come azione linguistica: aspetti sociali, epistemici e cognitivi* [Assertion as a Speech Action: Social, Epistemic and Cognitive Aspects]. Trieste, Edizioni Università di Trieste.

Labinaz, P. & M. Sbisà (2017): Credibilità e disseminazione di conoscenze nei social network. *Iride*, 30, 80, 63–86 DOI: <https://doi.org/10.1414/86522>.

Labinaz, P. & M. Sbisà (2021a): The Problem of Knowledge Dissemination in Social Network Discussions. *Journal of Pragmatics*, 175, 67–80 DOI: <https://doi.org/10.1016/j.pragma.2021.01.009>.

Labinaz, P. & M. Sbisà (2021b): Speech Acts and the Dissemination of Knowledge in Social Networks. In: Haberland, H., Yus, F. & C. Xie (eds.): *Approaches to Internet Pragmatics: Theory and Practice*. Amsterdam, John Benjamins, 145–172 DOI: <https://doi.org/10.1075/pbns.318.05lab>.

Sbisà, M. (1984): On Illocutionary Types. *Journal of Pragmatics*, 8, 93–112 DOI: 10.1016/0378-2166(84)90066-3. Reprinted in Marina Sbisà: *Essays on Speech Acts and Other Topics in Pragmatics*. Oxford, Oxford University Press, 23–42 DOI: <https://doi.org/10.1093/oso/9780192844125.003.0002>

Sbisà, M. (2007): How to Read Austin. *Pragmatics*, 17, 3, 461–473 DOI: 10.1075/prag.17.3.06sbi. In: Marina Sbisà: *Essays on Speech Acts and Other Topics in Pragmatics*. Oxford, Oxford University Press, 181–194 DOI: <https://doi.org/10.1093/oso/9780192844125.003.0011>.

Sbisà, M. (2020): Assertion Among the Speech Acts. In: Sandford, G. (ed.): *Oxford Handbook of Assertion*. Oxford, Oxford University Press, 159–178 DOI: 10.1093/oxfordhb/9780190675233.013.7. Reprinted in Marina Sbisà: *Essays on Speech Acts and Other Topics in Pragmatics*. Oxford, Oxford University Press, 267–287 DOI: <https://doi.org/10.1093/oso/9780192844125.003.0017>.

Sbisà, M. (2023): Introduction. In: Sbisà, M.: *Essays on Speech Acts and Other Topics in Pragmatics*. Oxford, Oxford University Press, 1–22 DOI: <https://doi.org/10.1093/oso/9780192844125.003.0001>.

Thi Nguyen, C. (2020): Echo Chambers and Epistemic Bubbles. *Episteme*, 17, 141–161 DOI: <https://doi.org/10.1017/epi.2018.32>.

Appendix: Example Sources

Acronym	Tweet Link	Date	Views	Comments	Reposts	Likes
T1	https://x.com/fanpage/status/1796927183153242130	1 June 2024	252739	669	630	874
T2	https://x.com/AUniversale/status/1796936825774202891	1 June 2024	89499	289	253	1366
T3	https://x.com/ri-tadallachiesa/status/1797020885959782698	1 June 2024	1300000	2276	777	5212

T4	https://x.com/fat-toquotidiano/status/1796917268430918072	1 June 2024	50361	385	190	482
T5	https://x.com/TELA-DOIOLANIUS/status/1797148754698453123	1 June 2024	37961	242	245	786